Identication of virtual grounds using virtual reality haptic shoes and sound synthesis

Serafin, Stefania; Turchet, Luca; Nordahl, Rolf; Dimitrov, Smilen; Berrezag, Amir; Hayward, Vincent

Published in:
Proceedings of the Eurohaptics 2010 Special Symposium

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Serafin, S., Turchet, L., Nordahl, R., Dimitrov, S., Berrezag, A., & Hayward, V. (2010). Identication of virtual grounds using virtual reality haptic shoes and sound synthesis. In A. Nijholt, E. O. Dijk, P. M.C. Lemmens, & S. Luitjens (Eds.), *Proceedings of the Eurohaptics 2010 Special Symposium: Haptic and Audio Visual Stimuli: Enhancing Experiences and Interaction* (1 ed., Vol. 10-01, pp. 61-70). University of Twente. CTIT Proceedings WP10-01 Vol. 10-01 No. 1

Identification of virtual grounds using virtual reality haptic shoes and sound synthesis

Stefania Serafin, Luca Turchet, Rolf Nordahl and Smilen Dimitrov
Medialogy, Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, Denmark
sts,tur,rn,sd@media.aau.dk

Amir Berrezag and Vincent Hayward
UPMC Univ Paris 06, Institut des Systèmes Intelligents et de Robotique,
4 place Jussieu,75005 Paris, France
amir.berrezag,vincent.hayward@isir.upmc.fr

Abstract

We describe a system which simulates in real-time the auditory and haptic sensation of walking on different surfaces. The system is based on physical models, that drive both the haptic and audio synthesizers, and a pair of shoes enhanced with sensors and actuators. In a discrimination experiment, subjects were asked to recognize the different simulated surfaces they were exposed to with uni-modal (auditory or haptic) and bi-modal (auditory and haptic) cues. Results show that subjects perform the recognition task better when using auditory feedback versus haptic feedback. The combination of auditory and haptic feedback only in some conditions significantly enhances recognition.

Keywords: walking simulation, physical models, audio-haptic interaction

1 INTRODUCTION

Research on multimodality has been focused on the interaction between vision and other modalities. However, lately the interest in investigating the interaction between other senses, such as touch and audition, has grown. This has been facilitated by the rapid development of haptic devices, together with efficient and accurate simulation algorithms. Several studies have indeed investigated audio-tactile phenomena such as the multimodal recognition of textures [10] and stiffness, both using physical stimuli [5] and simulations based on physical models [1].

The cited studies have focused on the interaction between touch and audition in hand-based interactions. To our knowledge, the interaction of auditory and haptic feedback in foot-based devices is still an unexplored topic. A notable exception is the work of Giordano et al., who showed that the feet were also effective at probing the world with discriminative touch, with and without access to auditory information. Their results suggested that integration of foot-haptic and auditory information does follow simple integration rules [8].

In previous research, we described a system able to recreate the auditory and haptic sensation of walking on different materials and presented the results of a preliminary surface recognition experiment [11]. This experiment was conducted under three different conditions: auditory feedback, haptic feedback, and both.

By presenting the stimuli to the participants passively sitting in a chair, we introduced a high degree of control on the stimulation. However, this method of delivery is highly contrived since it eliminates the tight sensorimotor coupling that is natural during walking and foot interaction. It is true for the auditory channel, but even more so for the haptic channel. In spite of these drastically constrained conditions, performance was surprisingly good.

In this paper, we extend such work by allowing subjects to walk in a controlled laboratory, where their steps are tracked and used to drive the simulation. We investigate whether introducing a higher level of interactivity will significantly enhance the recognition rates as well as the perceived quality and realism of the simulation.

2 Simulation Software

We developed a physically based synthesis engine able to simulate the auditory and haptic sensation of walking on different surfaces. Acoustic and vibrational signatures of locomotion are the result of more elementary physical interactions, including impacts, friction, or fracture events, between objects with certain material properties (hardness, density, etc.) and shapes. The decomposition of complex everyday sound phenomena in terms of more elementary ones has been an organizing idea in auditory display research during recent decades [7]. In our simulations, we draw a primary distinction between solid and aggregate ground surfaces, the latter being assumed to possess a granular structure, such as that of gravel.

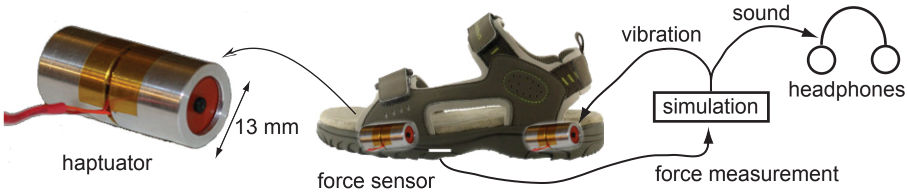

Figure 1: System (one shoe shown). Left: recoil-type actuation from Tactile Labs Inc. The moving parts are protected by an alumimum enclosure able to bear the weight of a person. Middle: approximate location of the actuators in the sandal. Right: system diagram showing the interconnections.

Figure 2: A picture of one pressure sensor and two actuators embedded in the shoes.

The impact model and its discretization are described elsewhere in detail [3]. The model has been recently adapted to the audio simulation of footsteps [12]. Here, we used the same model to drive the haptic and the audio synthesis. It is briefly recalled below.

A footstep sound may be considered to cause multiple micro-impacts between a sole, i.e., an *exciter*, and a floor, i.e., a *resonator*. Such interaction can be either discrete, as in the case of walking on a solid surface, or continuous, as in the case of a foot sliding across the floor.

In the simulation of discrete impacts, the excitation is brief and has an unbiased frequency response. The interaction is modelled by a Hunt-Crossley-type interaction where the force, f, between two bodies, combines hardening elasticity and a dissipation term [9]. Let x represent contact interpenetration and $\alpha > 1$ be a coefficient used to shape the nonlinear hardening, the special model form we used is

$$f(x, \dot{x}) = -kx^{\alpha} - \lambda x^{\alpha}\dot{x} \quad \text{if } x > 0, \quad 0 \text{ otherwise.}$$

The model described was discretized as proposed in [2].

If the interaction called for slip, we adopted a model where the relationship between relative velocity v of the bodies in contact and friction force f is governed by a differential equation rather than a static map [6]. Considering that friction results from a large number of microscopic damped elastic bonds with an average deflection z, a viscous term, $\sigma_2 v$, and a noise term, $\sigma_3 w$, to represent roughness, we have

$$f(z, \dot{z}, v, w) = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v + \sigma_3 w.$$

The force specified by these models is applied to a virtual mass which produces a displacement signal that is then processed by a linear shaping filter intended to represent the resonator.

Stochastic parameterization is employed to simulate particle interactions thereby avoiding to model each of many particles explicitely. Instead, the particles are assigned a probability to create an acoustic waveform. In the case of many particles, the interaction can be represented using a simple Poisson distribution, where the sound probability is constant at each time step, giving rise to an exponential probability weighting time between events.

We used this approach to model both solid and aggregate surfaces. A solid surface is represented by an impact and a slide. The impact model alone was used to recreate the sound and the feel produced when walking on wood. The friction model was tuned to simulate walking on creaking wood. To simulate walking on aggregate grounds, we used a physically informed sonic models (PhiSM) algorithm [4]. The synthesis was tuned to simulate snow and forest underbrush.

These algorithms were implemented as an extension to the Max/MSP platform[1] to drive both the auditory and haptic feedback.

3 Simulation Hardware

In order to provide both audio and haptic feedback, haptic shoes enhanced with pressure sensors have been developed. A pair of light-weight sandals was procured (Model Arpenaz-50, Decathlon, Villeneuve d'Ascq, France). This particular model has light, stiff foam soles that are easy to gouge and fashion. Four cavities were made in the tickness of the sole to accommodate four vibrotactile actuators (Haptuator, Tactile Labs Inc., Deux-Montagnes, Qc, Canada). These electromagnetic recoil-type actuators have an operational, linear bandwidth of 50–500 Hz and can provide up to 3 G of acceleration when connected to light loads. As indicated in Fig. 1 and Fig. 2, two actuators were placed under the heel of the wearer and the other two under the ball of the foot. There were bonded in place to ensure good transmission of the vibrations inside the soles. When activated, vibrations propagated far in the light, stiff foam. In the present configuration, the four actuators were driven by the same signal but could be activated separately to emphasize, for instance, the front or back activation, to strick a balance, or to realize other effects such as modulating different, back-front signals during heel-toe movements.

The sole has two FSR pressure sensors[2] whose aim was to detect the pressure force of the feet during the locomotion of a subject wearing the shoes. The two sensors were placed in correspondence to the heel and toe respectively in each shoe. The analogue values of each of these sensors were digitized by means of an Arduino Diecimila board[3] and were used to drive the audio and haptic synthesis.

A cable came out from each shoe, with the function of transporting the signals of the pressure sensors and of the actuators. Such cables were about 5 meters long, and they were connected through DB9 connectors to two 4TP (twisted pair) cables: one 4TP cable carried the sensor signals to a breakout board-which contained trimmers, that formed voltage dividers with the FSRs, which then interfaced to an Arduino board. The other 4TP cable carried the actuator signals from a pair of Pyle Pro PCA1[4] mini 2X15 W stereo amplifiers, driven by outputs from a FireFace 800 soundcard.[5] Each stereo amplifier handled 4 actuators found on a single shoe, and each output channel of the amplifier drove two actuators connected in parallel. The PC handled the Arduino through a USB connection, and the FireFace soundcard through a FireWire connection.

[1] www.cycling74.com
[2] I.E.E. SS-U-N-S-00039
[3] http://arduino.cc/
[4] http://www.pyleaudio.com/manuals/PCA1.pdf
[5] http://www.rme-audio.com/english/firewire/ff800.htm

4 CONTROLLING THE SYNTHESIS ENGINE

The audio-haptic synthesis engine is controlled by a ground reaction force (GRF), i.e., a force exerted by the ground on a body in contact with it. We first created a database of GRFs, by extracting the amplitude envelopes from different audio files of recorded footsteps, chosen among those available on the Hollywood Edge sound effects library.[6] Figure 3 shows the waveform of one of the footstep sounds used and its corresponding extracted GRF.

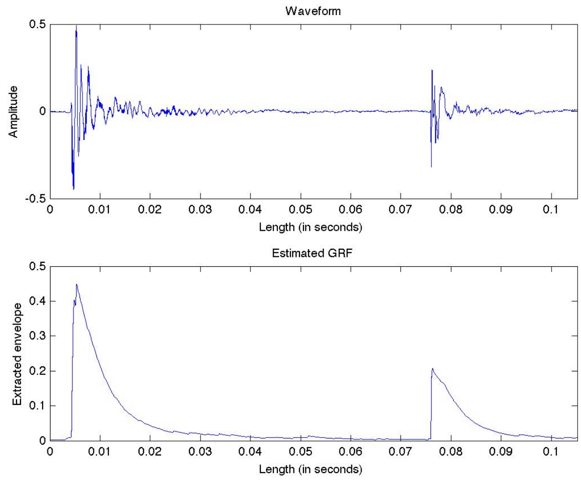

Figure 3: Time domain waveform of a footsteps sound (top) and its corresponding extracted GRF (bottom).

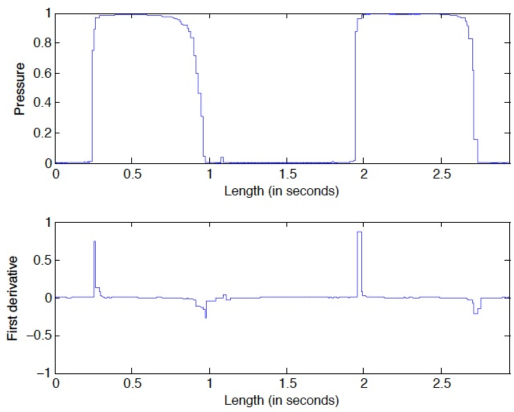

Figure 4: Waveform of a characteristic signal captured by the pressure sensors (top), and its derivative (bottom).

We then used the data captured by the pressure sensors and calculated their first time derivatives, which are related to the intensity with which the foot hits the ground. An example of data captured from the pressure sensors and its corresponding derivative is shown in Figure 4. The values of the first derivative were used as control for triggering the footsteps synthesizer, some GRFs corresponding to heel or toe

[6] www.hollywoodedge.com/

according to the activated sensor. Five types of heel and toes audio files were used for the purpose, and randomly chosen at the moment of triggering, giving rise to 25 possible combinations.

Each time the value of the first derivative became bigger than a threshold, the GRF corresponding to the activated sensor was triggered into the engine. More precisely we checked only positive changes in the derivative value, since we were interested in the generation of the sound when the step hits – and not when it leaves – the ground. Other thresholds, both on the signals and on their first derivatives, were used in order to handle some boundary conditions, like the standing of the subject, with the aim of controlling the generation of sound. Such thresholds were set in a phase of calibration of the system, which had to take into account the different weights of the subjects wearing the shoes, in order to have an average value suitable for all the possible cases.

5 EVALUATING THE SYSTEM

We conducted a between-subjects experiment whose goal was to investigate the ability of subjects to recognize the different simulated auditory and haptic stimuli they were exposed to. The experiment consisted of three conditions: in the first condition subjects were exposed to auditory feedback only, in the second condition to haptic feedback only, and in the third condition to a combination of audio and haptic feedback.

The sounds provided in the first condition were synthesized sounds generated in real time while subjects were walking using the interactive system described in the previous section.
The same stimuli were provided in the second condition by using the haptic shoes which provided haptic feedback in real-time. In the third condition, a combination of auditory and haptic stimuli was provided.

5.1 SETUP

All experiments were carried out in an acoustically isolated laboratory. The walking area was approximately 18 square meters, delimited by the walls of the laboratory. The setup for the first and third condition of the experiment consisted of the pair of sandals described in section 2, an Arduino board, a Fireface soundcard, a laptop and a set of headphones (Sennheiser HD 650, http://www.sennheiser.com).

In the first condition with audio only, the haptic actuators were not used, and the pressure sensors were driving the synthesis engine only to provide auditory feedback.
In the second condition (haptics only), participants were asked to wear earplugs and sound protection headsets instead of headphones, in order to block any audible feedback produced by the actuators. Indeed such sounds were not in the same quality range of the sounds designed to be conveyed through the headphones, and they could have biased the judgments of the participants.

In order to facilitate the navigation of the subjects, the wires coming out from the shoes in all setups, as well as the wires connecting the headphones to the soundcard, were linked to a bumbag or to snaplinks attached to trousers (see Fig. 5).

5.2 PARTICIPANTS

Thirty participants were divided in three groups (n = 10) to perform the between-subjects experiment. The three groups were composed respectively of 7 men and 3 women, aged between 20 and 35 (average age=24.6, sd=4.67), 9 men and 1 woman, aged between 20 and 31 (average age=23.4, sd=3.23), and 7 men and 3 women, aged between 21 and 25 (average age=22.7, sd=1.07). All participants reported normal hearing conditions. All participants were naive with respect to the experimental setup and to the purpose of the experiment.

The participants took in average about 10, 13 and 11 minutes for condition 1, 2 and 3 respectively.

5.3 TASK

During condition 1 and 3 the participants were asked to wear the pair of sandals and the headphones described in sections 2 and 5.1, and to walk in the laboratory. During condition 2 they were asked to wear the pair of sandals, earplugs and sound protection headsets, and to walk in the laboratory.

Participants were exposed to 16 trials in each condition. During the experiment, 8 stimuli were presented twice in randomized order. The stimuli consisted of audio and haptic simulations of footsteps sounds

Figure 5: A subject performing the experiment in the audio-haptic condition.

on the following surfaces: beach sand, gravel, snow (in particular deep snow), forest underbrush (a forest floor composed by dirt, leaves and branches breaking), dry leaves, wood, creaking wood and metal.

Participants were given a list of different surfaces to be held in one hand, presented as non-forced alternate choice. Such list included a range of materials wider than those presented in experiments.

During the act of walking they listened simultaneously to footsteps sounds and/or vibrations on a different surface according to the stimulus presented. The task consisted of answering by voice the following three questions after the presentation of the stimulus:

1. Which surface do you think you are walking on? For each stimulus choose an answer in the following list: 1) beach sand, 2) gravel, 3) dirt plus pebbles, 4) snow, 5) high grass, 6) forest underbrush, 7) dry leaves, 8) wood, 9) creaking wood, 10) metal, 11) carpet, 12) concrete, 13) frozen snow, 14) puddles, 5) water, 16) I don't know.

2. How close to real life is the sound in comparison with the surface you think it is? Evaluate the degree of realism on a scale from 1 to 7 (1=low realism, 7=high realism).

3. Evaluate the quality of the sound on a scale from 1 to 7 (1=low quality, 7=high quality).

The participants were informed that they could choose the same material more than one time and that they were not forced to choose all the materials in the list. In addition, they could use the interactive system as much as they wanted before giving an answer. When passed to the next stimulus they could not change the answer to the previous stimuli.

6 Results and discussions

The collected answers were analyzed and compared between the three conditions. Results are summarized in the confusion matrices reported in Table 1,2 and 3. Results confirm that auditory modality is dominant on the haptic modality and that the haptic task was more difficult than the other two.

Table 1: Confusion matrices with audio condition

	BS	GL	SW	UB	DL	WD	CW	MT	FS	CC	HG	DR	PD	WT	CP	—
BS	6	4	3					3			3				1	
GL		8		5	2						1	1	1	1		1
SW	1		15						3		1					
UB	1	4		6		1			6							2
DL	1	5		5	5				1		1					2
WD						10	2	1	1	2						4
CW							19									1
MT								13					1			6

Legend:	WD	wood	CW	creaking wood	SW	snow	UB	underbrush
	—	don't know	FS	Frozen snow	BS	beach sand	GL	Gravel
	MT	metal	HG	High grass	DL	dry leaves	CC	concrete
	DR	dirt	PD	puddles	WT	Water	CP	carpet

Table 2: Confusion matrices with haptic condition

	BS	GL	SW	UB	DL	WD	CW	MT	FS	CC	HG	DR	PD	WT	CP	—
BS	3	2	5	2			1		1			3				3
GL	5	1	3	1	1		1				2	2	1		1	2
SW	4		7		1		2		5							1
UB	1	3	5			2	1		1			3				4
DL	1	2	1		3				1			2			4	6
WD		2	2			6	3	1		1						5
CW			2			1	7		3	1	1			2		3
MT					1		5	7				2			1	4

Legend:	WD	wood	CW	creaking wood	SW	snow	UB	underbrush
	—	don't know	FS	Frozen snow	BS	beach sand	GL	Gravel
	MT	metal	HG	High grass	DL	dry leaves	CC	concrete
	DR	dirt	PD	puddles	WT	Water	CP	carpet

Table 3: Confusion matrices with audio-haptic condition

	BS	GL	SW	UB	DL	WD	CW	MT	FS	CC	HG	DR	PD	WT	CP	—
BS	8	4	4		1				2		1					
GL		18		2												
SW			13						6							1
UB		1	2	8	1				6			2				
DL	2	5		3	5		1		1		1	2				
WD						18				1						1
CW							20									
MT								14								6

Legend:								
	WD	wood	CW	creaking wood	SW	snow	UB	underbrush
	—	don’t know	FS	Frozen snow	BS	beach sand	GL	Gravel
	MT	metal	HG	High grass	DL	dry leaves	CC	concrete
	DR	dirt	PD	puddles	WT	Water	CP	carpet

Table 4: Average realism scores from a seven-point Likert scale.

	BS	GL	SW	UB	DL	WD	CW	MT
audio	5	5.1875	5.5	3.6667	5.1667	4.2	4.4211	3.1538
haptics	5	5	6	-	3.6667	4	4.1429	4.8571
combined	4.125	5.2778	5.4615	5.25	3.6	2.8333	3.95	3.0714

Table 5: Average quality scores from a seven-point Likert scale.

	BS	GL	SW	UB	DL	WD	CW	MT
audio	5.05	4.9722	5.275	4.7222	5.2222	4.5625	4.4211	4.5
haptics	3.8824	4.2222	4.7895	3.75	3.8571	3.8667	3.6875	3.9375
combined	4.5	5.15	5.6316	4.7	4.85	3.3684	3.7	4.0714

The addition of the haptic to the audio modality seemed to help the recognition only in few cases. Indeed percentages are on average slightly higher in condition 3 rather than condition 1, although an in-depth analysis shows significant difference only for gravel (χ^2 = 8.9011, df = 1, p-value = 0.00285) and wood (χ^2 = 5.8333, df = 1, p-value = 0.01573).

In the audio-haptic condition subjects recognized quite well most of the modeled surfaces. Very high percentages were found for gravel, wood, creaking wood.

An analysis performed on the wrong answers of conditions 1 and 3 reveals that on average subjects tended to classify erroneously a surface as another belonging to a same category (e.g., wood-creaking wood-concrete, snow-frozen snow, dry leaves-forest underbrush-dirt) rather than to different categories (e.g., wood-water, wood-gravel, metal-dry leaves). In particular considering the category snow-frozen snow, snow was recognized with a percentage of 90% in the audio condition and of 95% in the audio-haptic condition.
Moreover, what emerges from these results is the ability of the subjects in distinguishing materials in the same category for solid surfaces, and their difficulties in the recognition of aggregate surfaces (aspect also confirmed by the comments of the participants). The same tendencies were found in another experiment we conducted using the same physical models controlled by another system [12] as well as in a previous experiment with audio-haptic stimuli, where subjects were receiving the stimuli passively [11].

Subjects were able to distinguish at haptic level solid surfaces with a percentage of 53%, and were able to distinguish aggregate surfaces with a percentage of 74%.

As regards the percentage of the "I don't know" answers, it was higher for condition 2 (17.5%), rather condition 1 (10 %) and lower in condition 3 (5%) and this is an indication of the difficulty of the task proposed. Condition 2 took longer (13 minutes) rather than condition 1 (9 minutes and 30 seconds) and 3 (11 minutes). This is also an indication of the difficulty of the task proposed.

Table 4 and 5 report results on the perceived quality and realism of the simulation. The degree of realism was calculated by looking only at that data from correct answers, i.e., when the surfaces were correctly recognized. As far as the quality judgement is concerned, the data was based on all the answers different from "I don't know". The results are similar to those presented in [11], even if in the previous studies a lower number of actuators was present (two for each shoe instead of four) and the experiment was performed in passive conditions. Moreover, results show that the combination of auditory and haptic stimuli does not always enhance realism and perceived quality of the simulation.

7 Conclusion

In this paper, we introduced a real-time footsteps synthesizer able to provide audio and haptic feedback, and which is controlled by the user during the act of walking by means of shoes embedded with sensors and actuators. The system was tested in a between-subjects experiment.

Results show that subjects are able to recognize most of the synthesized surfaces using the interactive system with high accuracy. Similar accuracy can be noticed in the recognition of real recorded footsteps sounds, which is an indication of the success of the proposed algorithms and their control.

The developed system is ready to be integrated in computer games and interactive installations where a user can navigate.

In future work, we indeed plan to utilize the system in multimodal environments, and include visual feedback, to understand the role of the different sensorial modalities to enhance sense of immersion and presence in scenarios where walking plays an important role.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme under FET-Open grant agreement 222107 NIW - Natural Interactive Walking.[7]

[7] www.niwproject.eu

REFERENCES

[1] Avanzini, F. and Crosato, P. (2006). Haptic-auditory rendering and perception of contact stiffness. *Lecture Notes in Computer Science*, 4129:24.

[2] Avanzini, F. and Rocchesso, D. (2001). Modeling collision sounds: Non-linear contact force. In *Proc. COST-G6 Conf. Digital Audio Effects (DAFx-01)*, pages 61–66.

[3] Avanzini, F., Serafin, S., and Rocchesso, D. (2005). Interactive simulation of rigid body interaction with friction-induced sound generation. *IEEE transactions on speech and audio processing*, 13(5 Part 2):1073–1081.

[4] Cook, P. (1997). Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49.

[5] DiFranco, D., Beauregard, G. L., and Srinivasan, M. (1997). The effect of auditory cues on the haptic perception of stiffness in virtual environments. *Proceedings of the ASME Dynamic Systems and Control Division.*

[6] Dupont, P., Hayward, V., Armstrong, B., and Altpeter, F. (2002). Single state elastoplastic friction models. *IEEE Transactions on Automatic Control*, 47(5):787–792.

[7] Gaver, W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, 5(1):1–29.

[8] Giordano, B. L., Mcadams, S., Visell, Y., Cooperstock, J., Yao, H. Y., and Hayward, V. (2008). Non-visual identification of walking grounds. *Journal of the Acoustical Society of America*, 123(5):3412–3412.

[9] Hunt, K. H. and Crossley, F. R. E. (1975). Coefficient of restitution interpreted as damping in vibroimpact. *ASME Journal of Applied Mechanics*, 42(2):440–445.

[10] Lederman, S. (1979). Auditory texture perception. *Perception*.

[11] Nordahl, R., Berrezag, A., Dimitrov, S., Turchet, L., Hayward, V., and Serafin, S. (2010a). Preliminary experiment combining virtual reality haptic shoes and audio synthesis. In *Proc. Eurohaptics*.

[12] Nordahl, R., Serafin, S., and Turchet, L. (2010b). Sound synthesis and evaluation of interactive footsteps for virtual reality applications. In *Proc. IEEE VR 2010*.